QUADERNI

Università degli Studi di Siena
DIPARTIMENTO DI ECONOMIA POLITICA

TUSHAR KANTI NANDI

Parametric and Semiparametric Estimation of the Adoption of Work Teams

n. 484 - Luglio 2006

Abstract - This paper is concerned about the adoption of work teams and the factors that facilitate team adoption. It focuses on four factors - trade union, technological change, training of the workforce and shared mode of compensation. Both parametric and semiparametric estimation methods are used to estimate the association of these factors with team adoption. A nonparametric confidence band test is used to test the parametric specification of probit model. The test rejects the distributional assumption of the parametric probit model. The semiparametric estimates show that trade union density is not associated with team adoption while prodit sharing, new technology and training provisions for more employees facilitate the adoption of work teams.

Jel Classification: C14, C25, D23, J53
Keywords: Work Teams; Probit; Confidence Band; Semiparametric.

Tushar Kanti Nandi, Department of Economics, University of Siena

1 Introduction

The recent wave of renewed interest, among academics, business and policy communities, in the transformation of work organization has engendered a lively debate about the nature and implication of modern forms of industrial relations. The intensification of international competition has made many to believe that in order to gain a competitive edge in an ever changing business environment firms must change the way they organize the production of goods or services. The modern form of industrial relations is envisioned to be different from the traditional practice of hierarchical management system, narrowly defined job, little or no discretion on the part of workers. Rather, the new forms of capitalist production give greater autonomy to workers, place higher responsibility on lower level workers, giving them relevant decision making powers, and encourage cooperation instead of competition among the workers. And at the firm level this new system places greater thrust on quality rather than quantity of the goods or services it produces.

The observed departure from the traditional labour relation practices has been viewed by many economists as a move from rigid work rules to a new set of institutions that gives employers a higher level of flexibility in the organization of work. Work teams, quality circles and job rotation are used by the employers to encourage employee participation to achieve larger organizational goal. It is believed that these practices are the vehicles to bring together the management expertise, and workers' knowledge and creativity in the search for superior production routines. These alternative forms of work organization have received tremendous appreciation from some commentators as being able to result in higher productivity, better firm performance, and a happier and more productive labour force. Yet the application of these new forms of work organization is not wide spread, which is at odds with what one would predict considering the perceived beneficiary effects of applying them. So, it might be appropriate to ask why the performance enhancing approach has not been widely pursued, and what determines the adoption of such workplace practices. As to these questions, the theoretical views have been all but a unifying theme. Some view industry, technology and market condition to be determining factors of adoption of new forms of work organization. Yet others place more emphasis on the supporting compensation schemes that give appropriate incentive to the workers to perform in a changed environment. This paper analyzes the association of these factors to the adoption of work teams using establishment level data from UK. In particular we focus on four factors - trade union, profit sharing, technology and training - and analyze their association with the adoption of work teams.

The previous studies that address these issues are Osterman (1994) and Gittleman et al. (1998). This paper deviates from the earlier literature in two respects. One, they use an additive index of flexibility in work organization, constructed from the presence of different work organization practices like teams, job-rotation, quality-circle and others, as the dependent variable to estimate the association between different factors and the change in work organization. This method of constructing the index is more of a simplifying technique than a logical formulation stemming from a theory. Often economic theory can guide to construct an index that incorporates elements from several dimensions of organizational change - such as team based work organization is hypothesized to fit best with shared modes of compensation. However, an index of flexibility that incorporates elements along a single dimension can be misleading if there is complementarity or substitutability among the elements. Having said so, we also recognize the conceptual difficulty in constructing a suitable index of change in work organization, since it requires information about the implementation of work practices in specific contexts in order to understand the flexibility they generate to cope up with changes in other dimensions. However, though a broad picture emerges from this approach, in terms of association of the index to the factors hypothesized, conclusion about the effect of these factors on the adoption of specific work practices, that constitute the index, can not be directly drawn, since there is no reason to believe that the factors affect all the work practices in the same direction. When they do not, the estimate is convoluted by different opposing effects. In order to circumvent this problem, in this paper we focus on a single work organization practice, teams, to examine the association of different factors to its adoption. This allows us to draw unambiguous conclusion about the effects of the factors on the likelihood of adopting work teams. Second, the estimation method used commonly in the literature is a parametric model, either a logit or a probit, without any test of the model specification. We use probit and a semiparametric estimation method, in order to compare results from both methods. It turns out from the specification tests that probit is not the right specification of the data generating process. We find from semiparametric estimates that among the factors considered training is the main determinant of the adoption of work teams.

Rest of the paper is organized as follows. The second section reviews the theoretical views that have implication for the adoption of participatory practices or organizational change in general. The third section describes our data. The fourth section is concerned about estimation methods. We give brief description of parametric probit estimation method and the semiparametric method of Klein and Spady (1993). The fifth section gives estimation results and specification test for the parametric probit model. The last section concludes the paper.

2 Background Discussion

This section draws on several pieces of theory that help to understand the relationship between trade union, profit sharing, technology, and training on one side and the adoption of work teams or organizational change in general on other side.

2.1 Trade Union:

A theoretical perspective on how organizational changes occur can be found in evolutionary economics (Nelson and Winter 1982). The evolutionary perspective suggests that the established pattern of operation of a routine in an organization is often tacit and ingrained in the collective knowledge of the organization. Since organizational changes disrupt this collective knowledge, they are infrequent and subject to selection process. The organizational changes can occur in two ways – organization's search for superior routines and trial and error experimentation. In either way, the changes that result in positive gains for the organization are retained. The trail and error experimentation is often a chaotic and probabilistic process whose occurrence can not be easily amenable. And in the search process, organizations often look to the familiar technologies and organizational practices in the neighbourhood. In both the approaches, organization is subject to certain level of inertia (Pil and MacDuffie 1996).

Both with trial and error approach, and localized search, the adoption and the success of a new routine depend on the cooperation of the workers whose jobs are directly affected by the new routine. It is important for the management to secure workers' compliance, particularly that of trade union, for the adoption of new practices. Though on the management's part imitating the practices in the neighborhood can be justified as industry standard, frequent use of trial and error method can face union resistance because the changes disrupt the tacit and collective knowledge of the organization that ease the jobs of a larger section of workers. However, the relationship between the trade union and the change in work organization is not apparent. It is possible to find two equally prominent view on the relation, drawing from the debate on the effect of trade unionism on economic performance. The neoclassical school tends to focus on the monopoly view of trade unionism. The conceptual contents of this school is derived from the free functioning of market economy, where the market forces ensure the most efficient and productive utilization of labour. In this framework, the trade union is viewed as a source of imperfection, that keeps the market price of labour above the social opportunity costs. A corollary of this thesis is that in addition to keeping the wage rate higher than the market clearing rate, the trade union imposes restrictive work rules.

There are several reasons to cast doubt on this view. The neoclassical approach has at its core an abstract model of perfectly competitive economy, and the believe that the "first-best" outcome is attainable by removing the sources of imperfection. The proposition is limited to a static consideration of allocative efficiency. Nolan and Marginson (1990) note that in contrast to the static framework on which the neoclassical school draws, consideration of dynamic structure can shed light on the way the trade union plays a key role in shaping the economic outcome in one direction rather than another. Moreover, the neoclassical analysis of trade union is premised on the idea that the employment relationship will be characterized by shared interest, harmony and cooperation in the absence of trade union. This view is at odds with the findings of the classical writers that the employment relationship is characterized by the exchange of wages for human capacities and not some prespecified quantity of performed labour. Hence, there remains plenty of scope for conflict, opportunism over the amount, quality and range of tasks to be performed after the labour exchange has been consummated.

On the other hand, the institutional school stressed the voice function of the trade union. Moving from the static framework to a repeated interaction model, the proponents of institutional school suggest that by closing off the routes to increased profitability by means of wage-cutting and labour intensification the trade union can act as a spur to investment in new and more effective production systems (Freeman and Medoff 1984). This might lead to investment in human resources and associated change in work organization.

Another channel through which the union can affect the change in work organization is its effect on the workers' turnover rate. For firms, considering a switch over from the traditional hierarchical system to a participatory work organization, turnover is a major concern, since the frequent change in the composition of the participatory group can be costly and an impediment to achieve the desired outcome from the group. The trade union, by providing a voice to workers' discontent, can reduce the probability that individual workers exercise the exit option in case of discontent. Freeman (1980) shows that unionized organizations in the United States have lower quits and separations, and higher tenure than their non-unionized counter parts. Lucifora (1998) finds similar evidence from establishment level data of Italy. It is also likely that if the union can bargain a wage higher than that of alternative employment, turnover will be lower.

2.2 Training:

It is often argued that a trained workforce reduces the cost of changing the way work is organized within an organization. A better trained workforce gives the employer higher level of flexibility in assigning tasks to the employees, hence the opportunity to experiment with new routines. Apart from cost considerations, other functions of training can facilitate organizational change. The prime candidates are the effects on turnover and employee commitment. There is considerable empirical evidence that employee training reduces turnover. The employer concern for turnover is likely to be higher in participatory system than in a scientific management system. Job matching and screening arguments suggest that firms may use induction and training as a signalling mechanism to attract better quality workers. And continuous on-the-job training for a good worker/job match is provided so that training is associated with longer tenure (Barron et al. 1989). Firm specific training promotes organizational citizenship that is more important for the effectiveness of participatory work practices than for the centralized monitoring practices. Some form of training such as training to perform multiple tasks are complementary to the participatory work organization. Lynch and Black (1998) show that training is associated with the presence of high-performance work practices in U.S. establishments.

2.3 Compensation:

It is a common theme in the literature on modern manufacturing that employee participation and shared mode of compensation are complementary to each other (Milgrom and Roberts 1995, Holmstrom and Milgrom 1994). The argument is that the benefit of work teams are enhanced by having an incentive scheme that conditions individual compensation on the group performance. The theoretical debate on the issue has a long standing in economics. It represents two competing views - the earlier literature's focus on famous $\frac{1}{N}$ problem (Holmstrom 1982) and the recent literature's search for conditions where the problem is mitigated (Kandel and Lazear 1992, Che and Yoo 2002). The relevance of the free-riding problem is challenged by appealing to mutual monitoring and peer pressure aspects of participatory work practices and profit sharing scheme. However, the argument that the mutual monitoring function of profit sharing is an effective device to mitigate free-riding problems does not find much empirical support (Gaynor and Gertler 1995, Prendergast 1999). Che and Yoo (2002) show that the effectiveness of profit sharing depends on the complementarity of the effort choices of the team members. The effectiveness derives from the increased ability to punish in an infinitely repeated game. If the effort levels are complementary, as is likely to be the case in participatory work practices, the incentive effect of profit sharing more than offset the free-riding effect.

2.4 Technology:

The technology dependence of team adoption is pronounced by several economists. Alchian and Demsetz (1972) maintain that team production arises when the output is not separable in contributions of individual inputs or when the total output from cooperative inputs is high enough than the sum of separable outputs to cover the cost of monitoring and disciplining the input behavior. The technology of an organization often defines the set of feasible work organization practices. In a Tayloristic technological set up, it is often argued, most of the jobs are routine works and there is little opportunity on the part of workers to put discretionary efforts. Hence, the participatory work practices are less likely to be adopted in such a set up. On the other hand, the post Taylotistic production methods exploits both the management expertise and workers knowledge of the jobs they perform. Hence, the technological change from the former to latter is more likely to be associated with the adoption of new work organization practices. The new technology may constitute an example of what Pil and MacDuffe (1996) refer to as an "unfreezing" of the existing way of doing business, that is, a disruption that moves the organization away from the status quo. At such times, the cost of experimenting may be lower, since the organization is likely to be in a state of flux away.

Table 1: Variable definition and descriptives

Variable	Definition	Mean	Std. Dev.
Team	1 if almost all (80% or more) employees of largest occupational group work in teams, 0 otherwise	0.5221	0.4998
Size	Number of employees	322.7773	1133.6530
Union Density	Number of union members / Size	26.0216	33.4546
Profit Pay	1 if profit sharing scheme for largest occupational group 0 otherwise	0.4979	0.5003
New Technology	1 if new technology in last 12 months 0 otherwise	0.8372	0.3694
Training (Ref. no training)			
Training 1-19	1 if 1 to 19% employees had off-the-job training, 0 otherwise	0.2237	0.4170
Training 20-59	1 if 20 to 59% employees had off-the-job training, 0 otherwise	0.2773	0.4479
Training 60-100	1 if 60 or more employees had off-the-job training, 0 otherwise	0.3613	0.4806
Part-time	Numebr of part-time employees / Size	0.2151	0.2846
Female	Number of female employees / Size	0.4060	0.2604
Industry (Ref. manufacturing)			
EG&W	1 if the workplace in Elecetricity, Gas & Water sector, 0 Otherwise	0.0651	0.2469
Construction	1 if in Construction sector, 0 otherwise	0.0525	0.2232
Wholesale & Retail	1 if in Wholesale & Retail sector, 0 otherwise	0.2489	0.4326
Hotel & Restaurant	1 if in Hotel & Restaurant sector, 0 otherwise	0.0819	0.2744
Transport & Communication	1 if in Transport & Communication sector, 0 otherwise	0.0735	0.2611
Financial Services	1 if in Financial Service sector, 0 otherwise	0.0714	0.2577
Other Business Services	1 if in Other Business Service sector, 0 otherwise	0.1460	0.3533

3 Data

Data for this study come from the British Workplace Employee Relations Survey (WERS), 1998. The Workplace Employee Relations Survey[1] contains detailed information about labour relations, work organization, compensation and training provision of 2191 British workplaces with 10 or more employees. Only the private sector workplaces with 25 or more employees are considered for the present study.

The management questionnaire of the survey asks what percentage of the workers of the largest occupational group work in formally designated teams. Answer to this question is allowed to be in several intervals ranging from 0% to 100%. We construct a binary variable that takes value 1 if almost all (80% to 100%) of the employees in largest occupational group work in teams

[1]Department of Trade and Industry, Advisory, Conciliation and Arbitration Service, Workplace Employee Relations Survey : Cross-Section, 1998 [computer file]. 6th ed. Colchester, Essex: UK Data Archive [distributor], 23 January 2001. SN: 3955.

in the establishment, 0 otherwise. In the empirical analysis that follows we test several hypothesis regarding adoption of work teams.

The independent variables correspond to the factors discussed in the previous section, namely - trade union, technological change, training provision and compensation structure. In order to capture trade union effect we use the trade union density in the establishment as control. We also control for the size and industry effects in the estimation. Observations with missing values are omitted, which leads to a valid data set of 952 establishments in eight industries. Table 1 defines the variables and gives summary statistics.

There are 52% establishments that have work teams. Trade union membership is little higher in establishments with teams (26.9%) than in establishments without teams (25.1%). With regard to profit sharing scheme, teams are less prevalent in establishments with no such scheme (44%) than in those with (60.6%). Similar is the pattern for new technology. The correlation between teams and off-the-job training is negative for training provision for less than 20% employees. The correlation becomes positive for training for higher percentage of employees.

4 Estimation

Since the dependent variable we want to explain - presence of work teams - takes only two values (0 and 1), we can express the model as

$$\begin{aligned} y &= 1 \text{ if } v\left(x,\theta\right) \geq u \\ &= 0 \text{ otherwise,} \end{aligned} \tag{1}$$

where $v\left(.\right)$ is a known function (usually assumed to be equal to $x\theta$), x is a row vector of exogenous variables, θ a column vector of unknown parameters, and u a random disturbance. Let $F(u|x)$ denote the cumulative distribution of u conditional on x. Then the probability of having teams is

$$P(y=1|x,\theta) \equiv \Pr(u < v(x,\theta)|x) = F[v(x,\theta)|x] \tag{2}$$

A standard choice for estimation of this model is to use a probit specification. The probit model belongs to a parametric family that assumes the functional form of the distribution of the random component of the model. In the case of probit, the assumption is that the distribution of the error term is normal. The estimates are obtained by maximizing a likelihood (usually log likelihood) function that depends on the parameters of the model. One drawback of the parametric estimation is that inconsistent estimates are obtained if the distribution of the error term is misspecified. This leads to a search for estimation methods that do not require specifying the error distribution.

There already exist several semiparametric estimation methods that do not require distributional assumption. Cosslett (1983) develops a *distribution free* consistent estimator. But, this estimator requires that the distribution of x is known up to finite-dimensional parameters. The approach here is to choose the unknown distribution function jointly with θ to maximize the likelihood function. Ichimura (1993), and Klein and Spady (1993) develop $N^{1/2}$ consistent, asymptotically normal semiparametric estimators of θ. These estimators permit certain forms of heteroskedasticity of u. The Klein and Spady estimator also achieves the asymptotic efficiency bound of Cosslett (1987). Horowitz (1992) provides a smoothed version of maximum score estimator, originally developed by Manski (1975). The smoothed maximum score estimation allows heteroskedasticity of unknown form, but requires the assumption that the median of $(u|x)$ is zero. The convergence rate of this estimator is between $N^{1/3}$ and $N^{1/5}$, which can be made arbitrarily close to $N^{1/2}$, but that requires rather restrictive assumptions.

In this paper, we consider two estimation methods: probit and semiparametric estimation proposed by Klein and Spady (1993). It is important to note that the Klein and Spady (K-S) estimator has most of the features of parametric estimators, namely $N^{1/2}$ consistency, asymptotic normality and efficiency. It also serves to compare the estimation results with those obtained from the parametric probit model. The following subsections give a brief description of the probit maximum likelihood method and the semiparametric pseudo maximum likelihood method of K-S.

4.1 Probit maximum likelihood

The probit model assumes a normal distribution for F. The maximum likelihood (ML) estimator of θ is obtained by maximizing the log likelihood function

$$LL(\theta) = \sum_{i=1}^{n} y_i \ln\left[F(v(x,\theta)|x)\right] + (1-y_i)\ln\left[1 - F(v(x,\theta)|x)\right] \qquad (3)$$

with

$$F(v(x,\theta)|x) = F(x,\theta) = \Phi(v(x,\theta)/\sigma_u)$$

where Φ is the cumulative normal distribution function and σ_u^2 the variance of u. The probit model is identified only up to a constant scale factor. It is standard to achieve the scale normalization by setting σ_u^2 equal to 1.

4.2 Pseudo maximum likelihood

When F is unknown, Cosslett (1983) proposes to choose F jointly with θ to maximize the likelihood function in (3). Though the resulting estimator is consistent, once the distribution function is replaced by its ML estimate, the resulting concentrated likelihood is not a smooth function of θ. Consequently, it becomes difficult to establish the asymptotic distribution for the estimator of θ. K-S propose to select θ so as to maximize a semiparametric likelihood that is a smooth function of θ and that locally (for θ in the neighborhood of θ_0, the true value) approximates the corresponding parametric likelihood.

Suppose that the model satisfies the single index restriction, i.e. $E[y|x] = E[y|v(x,\theta)]$. Then the quasi-likelihood function that approximates the parametric likelihood can be obtained by replacing $P[y=1|v(x,\theta)]$ with a tractable function $P[v(x,\theta),\theta]$ that locally approximate it. $P[v(x,\theta),\theta]$ can be written as

$$\begin{aligned} P[v(x,\theta),\theta] &\equiv \Pr[u < v(x,\theta)|v(x,\theta)] \\ &= \Pr[u < v(x,\theta)].g_{v|u<v(x,\theta)}(v(x,\theta))/g_v(v(x,\theta)) \end{aligned} \tag{4}$$

where $\Pr[u < v(x,\theta)]$ is the unconditional probability of $u < v(x,\theta)$, $g_{v|u<v(x,\theta)}(v(x,\theta))$ the density of $v(x,\theta)$ conditional on $u < v(x,\theta)$ and $g_v(v(x,\theta))$ the unconditional density of $v(x,\theta)$. It is easy to see that at the true value of the parameter, θ_0,

$$\Pr[u < v(x,\theta)] = \Pr[y=1]$$

and

$$g_{v|u<v(x,\theta)}(v(x,\theta) = g_{v|y=1}(v(x,\theta)).$$

So, an estimate for (4) can be obtained by replacing $\Pr[y=1]$ by the sample proportion of observations with $y=1$ and both the densities (conditional and unconditional) by their nonparametric estimates. A desirable feature of this formulation is that at $\theta=\theta_0$, and under single index restriction, $P[v(x,\theta),\theta]$ is equivalent to the true probability, i.e.

$$P[v(x,\theta),\theta] = \Pr[u_0 < v(x,\theta_0)|v(x,\theta_0)] \equiv \Pr[y=1|v(x,\theta_0)] = \Pr[y=1|x].$$

In application, $P_i[v_i(x,\theta),\theta]$ can be estimated as follows. Let,

$$g_{yv}(v_i(x,\theta),\theta) \equiv \Pr(y).g_{v|y}(v_i(x,\theta),\theta), \qquad (y=0,1)$$

and

$$g_v \equiv g_{1v} + g_{0v}$$

From (4),

$$P_i(v_i,\theta) = \frac{g_{1v}(v_i,\theta)}{g_v(v_i,\theta)} \tag{5}$$

Now, it remains to estimate $g_{yv}(v_i,\theta)$ for $y = 0,1$. K-S propose to estimate $g_{yv}(.)$ with kernel function that is symmetric, integrates to one and has bounded second moment. Kernel can be either bias-reducing or adaptive with local smoothing. For a bias-reducing kernel $K(z)$, with $\int z^2 K(z)dz = 0$, estimate of $g_{yv}(.)$ is given by,

$$\hat{g}_{yv}(v_i(x,\theta)) = \frac{1}{n-1}\sum_{j\neq i}^{n} \frac{1\{y_j = y\}}{h_n} K\left[\frac{v_i - v_j}{h_n}\right], \qquad (y=0,1) \tag{6}$$

where n is the number of observations and bandwidth h_n is such that $n^{-1/6} \prec h_n \prec n^{-1/8}$.

In order to ensure uniform convergence, an adjustment factor is used to guard against small estimated densities. The adjustment factor takes the following form

$$\widetilde{\delta}_{yn} \equiv h_n^a[e^z/(1+e^z)], \qquad z \equiv [(h_n^b - \hat{g}_{yv}(v,\theta))/h_n^c], \qquad (y=0,1).$$

and

$$\widetilde{\delta}_n \equiv \widetilde{\delta}_{0y} + \widetilde{\delta}_{1y}\ .$$

where $0 < b < c$ and $a > 2b + 2c > 0$. With these adjustment factors the estimated probability function is

$$\widetilde{P}(v,\theta) = \frac{\hat{g}_{1v}(v,\theta) + \widetilde{\delta}_{1n}(v,\theta)}{\hat{g}_v(v,\theta) + \widetilde{\delta}_n(v,\theta)} \tag{7}$$

The adjustment factors control the rate at which the numerator and denominator of estimated probability function tend to zero. For small estimated densities the adjustment factors behave exponentially like h_n^a. For sufficiently large estimated densities, the adjustment factors tend exponentially to zero as no adjustment is required.

Finally, with an appropriate likelihood trimming function, $\widetilde{\tau}(.)$, the K-S estimates are obtained by maximizing the following log likelihood function

$$Q(\widetilde{P}(\theta)) = \sum_{i=1}^{n}(\widetilde{\tau}_i/2)\left(y_i \ln[\widetilde{p}_i(\theta)^2] + (1-y_i)\ln[(1-\widetilde{P}_i(\theta))^2]\right) \qquad (8)$$

The trimming sequence is introduced for the sake of the normality argument. This ensures that the estimated probabilities converges to the true probabilities at a sufficiently fast rate. In application, the trimming sequence can be omitted, and the resulting estimates will still be consistent. The arguments of the log function have been squared to guard against the fact that with a bias reducing kernel the estimated densities, hence the estimated probabilities, can be negative. Since this might happen in a finite sample application, squaring the arguments of the log function ensures that the likelihood function remains well behaved even in this case.

5 Results and Tests

This section gives our estimation and test results. For all estimation we use the log of the number employees in a workplace for its size. Table 2 gives the probit estimates, with standard errors corresponding to the covariance matrix that is computed from the Hessian and the cross product of the first derivatives of the likelihood function[2].

The estimated coefficients suggest insignificant effect of union density on the adoption of work teams. Profit sharing scheme appears to have significant effects on the adoption of work teams. The coefficient for new technology is not significant. Training for more employees is positively associated with the adoption of work teams. Before interpreting the results, we turn to a specification test of probit model based on a nonparametric confidence band.

The probit specification can be tested with a nonparametric estimation of the confidence band. Let θ_p and $\sum_p$ denote the maximum likelihood estimates of the parameters and the covariance matrix, respectively, of the probit model. Let, $V(x) = x\sum_p x'$. If the probit model is the correct description of the data generating process, a nonparametric regression of y on $x\theta_p/V(x)^{1/2}$ yields a consistent estimate of the cumulative normal distribution function,

[2]The covariance matrix, $\sum_\theta = \left(\frac{1}{n}\sum_{i=1}^{n}\frac{\delta^2 L}{\delta\theta\delta\theta'}\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}(\frac{\delta L}{\delta\theta})'(\frac{\delta L}{\delta\theta})\right)\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\delta^2 L}{\delta\theta\delta\theta'}\right)^{-1}$

Table 2: Probit estimates

	Coefficient	S. E.	T-statistics
Intercept	-0.9963	0.2608	-3.82
Size	0.0050	0.0427	0.12
Union Density	-0.0003	0.0016	-0.21
Profit Pay	0.2519	0.0902	2.79
New Technology	0.1995	0.1173	1.70
Training 1-19	0.1744	0.1496	1.17
Training 20-59	0.5683	0.1473	3.86
Training 60-100	0.8044	0.1463	5.50
Part-time	0.0235	0.2452	0.10
Female	0.4665	0.2526	1.85
EG&W	0.3191	0.1955	1.63
Construction	0.1070	0.2063	0.52
Wholesale & Retail	0.0621	0.1542	0.40
Hotel & Restaurant	-0.2618	0.2002	-1.31
Transport & Communication	-0.1052	0.1804	-0.58
Financial Services	0.4093	0.1979	2.07
Other Business Services	0.1653	0.1525	1.08

$\Phi(.)$. The statistical significance of the difference between the nonparametric regression curve and $\Phi(.)$ can be assessed by constructing a uniform confidence band for the nonparametric regression function (Horowitz 1993). The following *proposition* of Horowitz (1993) outlines the construction of a nonparametric confidence band for the probit model:

> Let $\widetilde{F}$ denote the kernel regression of y on $x\theta_p/V(x)^{1/2}$. Let the kernel, K, be a probability density that is symmetrical about 0, has bounded support, and whose first derivative has bounded variation. Let the bandwidth be $w_n = h_n^\lambda$, where $1 < \lambda < 5/3$ and $h_n \propto n^{-1/5}$. Let $g(.)$ denote the probability density function of $x\theta/V(x)^{1/2}$. Let $\hat{g}$ denote the kernel estimate of g based on $x\theta/V(x)^{1/2}$, kernel K, and bandwidth w_n. Let S be a closed interval of the real line on which g is strictly positive. Assume that $P(y = 1|x) = \Phi(x\theta/V(x)^{1/2})$ and that g is twice differentiable. Then for any real z,

$$\lim_{n\to\infty} P[(0.4\lambda \ln n)^{1/2}\{(nw_n/c)^{1/2}\sup_{v\in S}[\hat{g}(v)/\sigma_n^2(v)]^{1/2}$$
$$|\widetilde{F}(v) - \Phi(v)| - d_n\} < z] = \exp(-2\exp(-z))$$

where

$$\sigma_n^2(v) = \widetilde{F}(v)[1 - \widetilde{F}(v)]$$

$$d_n = (0.4\lambda \ln n)^{1/2} + (0.4\lambda \ln n)^{1/2} \ln[c^*/(2\pi^2)]^{1/2},$$

$$c = \int K(u)du$$

$$c^* = (2c^{-1}) \int K'(u)^2 du$$

From this *proposition* the confidence band can be constructed as

$$Confidence\ band = \widetilde{F}(v) \pm \left\{ \left[\frac{z}{(0.4\lambda \ln n)^{1/2}} + d_n \right]^{1/2} \left[\frac{\sigma_n^2(v)c}{nw_n\hat{g}} \right]^{1/2} \right\}$$

where z is such that $\exp(-2\exp(z)) = 1-\alpha$, α being the level of significance.

Figure 1 shows the 95% confidence band along with the cumulative normal distribution. We use a Gaussian kernel, truncated at -100 and 100, since compact support is required by the *proposition*. Horowitz (1993) notes that Gaussian kernel can be used if the truncation point is made sufficiently large. The figure shows that a large part the standard normal distribution lies outside the nonparametric confidence band. This indicates that the probit model is not the right specification of the data generating process.

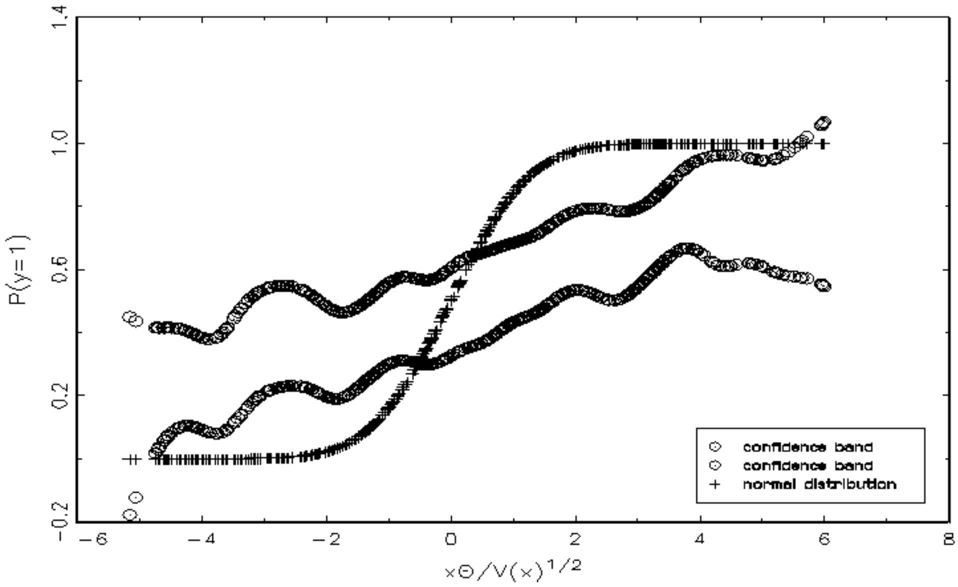
The semiparametric estimates are given in table 3. The estimates are obtained by using 7th order Gaussian kernel[3] and bandwidth equal to $n^{-1/7}$.

[3]The 7th order Gaussian kernel takes the form

$$K_7(w) = \frac{1}{8}(105 - 105w^2 + 21w^4 - w^6)\phi(w)$$

where $w = \frac{x - x_i}{h}$ and $\phi(.)$ is standard normal density.

Figure 1: Figure 1: Nonparametric Confidence bands



We do not use likelihood trimming sequence, as they are not necessary for the consistency of the estimates. Intercept is not included since it is not identified in K-S estimation method, and normalization is achieved by setting the coefficient of size equal to one.

The semiparametric estimates give a slightly different picture as compared to that of probit estimates. Trade union density is insignificant. Profit sharing scheme has significant association with team adoption. The coefficient of new technology is significant and positive. Training has positive association with team adoption, as compared to the omitted group - no training. The estimates of the training variables shows that the association is stronger the wider is the provision of training.

Our results show that establishments wide spread employee training facilitate the adoption of work teams. The asset based theory of firms conjectures that competitiveness of firms in global economy is more dependent on quality of it's workforce than on other easily imitable factor such as technology. The theory emphasizes that since competitive gains resulting from competence and distictiveness of the workforce can not be imitated, the workforce oriented competitive edge gives firms a unique way to achieve business success. The workforce oriented competitive strategy requires that employees work in

Table 3: Semiparametric estimates

	Coefficient	S. E.	T-statistics
Size	1.0000	.	.
Union Density	-0.0006	0.0035	-0.18
Profit Pay	0.8598	0.2144	4.01
New Technology	0.4965	0.0524	9.47
Training 1-19	-0.0144	0.0789	-0.18
Training 20-59	0.7934	0.0512	15.49
Training 60-100	1.9012	0.1887	10.07
Part-time	0.6469	0.2963	2.18
Female	0.6027	0.1607	3.75
EG&W	0.4670	0.0919	5.08
Construction	-0.2790	0.1278	-2.18
Wholesale & Retail	0.0794	0.0609	1.30
Hotel & Restaurant	-0.1091	0.0722	-1.51
Transport & Communication	-0.8747	0.1405	-6.23
Financial Services	-0.2315	0.2946	-0.79
Other Business Services	-0.1574	0.0616	-2.56

an environment that promote their discretionary effort and creativity. The training for the workforce plays a important role in shaping their practical and creative knowledge in a way that obtains organizational goals. Our finding that training is the most important factor associated with work teams is consistent with the asset based theory of firms.

These results are not readily comparable to other studies on the same issue, for two reasons (as has been mentioned before): others use a composite index of organizational change and the estimation methods employed by them have invariably been parametric, without any specification test. However, the union effects in Gittleman et al. (1998) and in Osterman (1994) are not significant. The training effect in our paper is consistent with their findings. The industry dummies in Gittleman et al. (1998) are negative significant (with respect to the omitted group, manufacturing). In our results we find that though establishments in some sectors (Construction, Transport & Communication, and Other Business Services) are less likely to adopt work teams, those in Electricity, Gas & Water sector are more likely to adopt work teams compared to manufacturing sector.

6 Conclusion

This paper addresses the issue of adoption of a commonly used form of work organization - work teams. It tests the empirical relevance of several pieces of theory that relates team adoption to trade unionism, technological change, training provision and shared mode of compensation. Departing from the common practice of parametric estimation, we estimate the team adoption using both parametric and semiparametric specification. The nonparametric confidence band test rejects the normality assumption of the probit model. However, the probit estimates are qualitatively similar to the semiparametric estimates except for profit sharing.

The semiparametric estimates show that trade union density is not associated to team adoption. The largely believed implication of complementarity (between work teams and share mode of compensation) hypothesis is supported by our findings. However, in line with others, we find that adoption of new technology and more training is associated with adoption of work teams. It is evident that establishments in the manufacturing industry are more likely to have work teams than those in other industries. However, our findings show that the establishments in Electricity, Gas & Water sector are more likely to have teams than manufacturing establishments.

References

[1] **Alchian, A. A. and Demsetz, H. (1972)**, Production, Information Costs, and Economic Organization. *American Economic Review* 62, 777-795.

[2] **Barron, J. M., Black, D. A. and Loewenstein, M. A. (1989)**. Job matching and on-the-job training. *Journal of Labor Economics* 7, 1-19.

[3] **Che, Y. and Yoo, S. (2002)**. Optimal incentives for teams. *American Economic Review* 91, 525-541.

[4] **Cosslett, S. R. (1983)**, Distribution-free maximum likelihood estimation of binary choice model. *Econometrica* 51, 765-782.

[5] **Cosslett, S. R.(1987)**, Efficiency bounds for distribution-free estimators of binary choice and censored regression models. *Econometrica* 55, 551-586.

[6] **Freeman, R. B. (1980)**. The exit-voice tradeoff in the labor market: Unionism, job tenure, quits, and separations. *Quarterly Journal of Economics* 94, 643-673.

[7] **Freeman, R. B. and Medoff, H. (1984)**. *What do unions do?* New York: Basic Books.

[8] **Gaynor, M. and Gertler, P. J. (1995)**. Moral hazard and risk spreading in partnerships. *Rand Journal of Economics* 26, 591-613.

[9] **Gittleman, M., Horrigan, M. and Joyce, M. (1998)**, "Flexible" workplace practices: Evidence from nationally representative survey. *Industrial and Labor Relations Review* 52, 99-115.

[10] **Holmstrom, B. (1982)**. Moral hazard in teams. *Bell Journal of Economics* 13, 324-340.

[11] **Holmstrom, B. and Milgrom, P. (1994)**. The firm as an incentive system. *American Economic Review* 84, 972-991.

[12] **Horowitz, J. L. (1993)**, Semiparametric estimation of a work-trip mode choice model. *Journal of Econometrics* 58, 49-70.

[13] **Horowitz, J. L. (1992)**, A smoothed maximum score estimator for the binary response model. *Econometrica* 60, 505-531.

[14] **Ichimura, H. (1993)**, Semiparametric least squares (SLS) and weighted SLS estimation of single-index models. *Journal of Econometrics* 58, 71-120.

[15] **Kandel, E. and Lazear, E. P. (1992)**. Peer pressure and partnerships. *Journal of Political Economy* 100, 801-817.

[16] **Klein, R. W. and Spady, R. H. (1993)**, An efficient semiparametric estimator for binary response model. *Econometrica* 61, 387-421.

[17] **Lucifora, C. (1998)**, The impact of unions on labour turnover in Italy: evidence from establishment level data. *International Journal of Industrial Organization* 16, 353-376.

[18] **Lynch, L. M. and Black, S. E. (1998)**. Beyond the incidence of employer-provided training. *Industrial and Labor Relations Review* 52, 64-81.

[19] **Manski, C. (1975)**, Maximum score estimation of the stochastic utility model of choice. *Journal of Econometrics* 3, 205-228.

[20] **Milgrom, P. and Roberts, J. (1995)**. Complimentarities and fit: Strategy, structure and organizational change in manufacturing. *Journal of Accounting and Economics* 19, 179-208.

[21] **Nelson, R. R. and Winter, S. (1982).** *An evolutionary theory of economic change.* Harvard University Press.

[22] **Nolan, P. and Marginson, P. (1990)**, Skating on thin ice? David Metcalf on trade unions and productivity. *British Journal of Industrial Relations* 28, 227-247.

[23] **Osterman, P. (1994)**. How common is workplace transformation and who adopts it? *Industrial and Labor Relations Review* 47, 175-188.

[24] **Pil, F. K. and MacDuffie, J. P. (1996)**. The Adoption of high involvement work practices. *Industrial Relations* 35, 423-455.

[25] **Prendergast, C. (1999)**. The provision of incentives in firms. *Journal of Economic Literature* 37, 7-63.